

AN APPLICATION OF PEER ASSESSMENT IN HIGHER EDUCATION

Sami ŞAHİN
Gazi University

ABSTRACT

The purpose of this study is to practice peer evaluation and to determine if the result of the evaluation shows similarity with the lecturer evaluation, thus to make assumption about the validity of peer evaluation in higher education. For this purpose, students of “Specific Teaching Methods I” class, which is included in the 3. Class of Department of Computer and Instructional Technologies Teaching, Gazi Faculty of Education, Gazi University, evaluated peers for their term project. In order to determine the similarity between peer and lecturer evaluation, the correlation between scores of lecturer and scores gained by peer evaluation is calculated. The result of the study revealed that peer evaluation showed similarity with lecturer evaluation.

KEY WORDS: Higher Education, Peer Evaluation, Validity

INTRODUCTION

In such a world, in which knowledge is progressively increasing and thus constructivist approaches are progressively becoming widespread; educators feel, and also researchers know, that classical evaluation methods fall behind in evaluation of active learning. The effect of this need forces educators to use active evaluation methods and also orients researchers to this area. Portfolios, performance evaluation, peer and self-evaluation are some of known alternative evaluation methods. However, in order these methods to be recognized and used effectively by educators, it is important to introduce the guiding principles and maintain high validity during implementation of these active evaluation methods.

As an alternative assessment method, peer evaluation is also utilized as a learning activity in addition to assess learning performance (Freeman, 1995). However, lecturers and students have some doubts about the validity level of evaluation, made by students, at the same knowledge level, about each other (Holroyd, 2000). This attitude seems to be an important barrier in front of use of peer evaluation. There is need for researches, revealing that, when it is implemented in a right way, peer evaluation demonstrates results, which are comparable with lecturer evaluation (Fry, 1990).

In this study, validity level of peer evaluation, when compared to lecturer evaluation, has been researched. For this purpose, students prepared a term paper for an end-term evaluation in 2007 Spring Term of “Special Teaching Methods I” class, which is included in the 3. Class of Department of Computer and Instructional Technologies Teaching, Gazi Faculty of Education, Gazi University and made peer evaluation. In order to determine the validity of peer evaluation, lecturer scores were taken as criteria and the correlation between lecturer scores and scores gained by peer evaluation was examined.

CONCEPTIONAL FRAMEWORK

In this chapter, some theoretical concepts, considered important in the scope of the study, have defined and basic features have been identified.

Peer Evaluation

Peer evaluation, considered as an alternative evaluation method, which includes both learning and evaluation, progressively attracts attention in higher education (Falchikov, 2001). This constitutes a parallelism with such recent improvements as cooperative learning, which attracts attention in higher education. Peer evaluation is described as an evaluation method, in which individuals evaluate each other according to certain criteria (Falchikov, 1995). In this evaluation method, students evaluate the quality of studies of their class-mates and provide feedback to each other (Van den Berg, Admiraal, and Pilot, 2006)

Advantages and Limitations of Peer Evaluation

Students interpret the studies of peers, while evaluating them, and thus they contribute to their own learning (Topping, 1998; Fallows, and Chandramohan, 2001). The advantages of peer evaluation in addition to increasing motivation of students for learning, are as follows: taking the responsibility of their own learning, making evaluation a part of learning, considering mistakes not as failure but as opportunity for re-learning, putting into practice the skills for knowledge transfer, using peer evaluation as a self-evaluation form, providing deep-learning instead of superficial learning (Brown, 1998). In addition to these, researchers indicate some important features of peer evaluation as follows (Zariski 1996; Van den Berg et. al., 2006):

- Evaluation is a part of learning
- Students become more autonomous, responsible and participative
- Students not only monitor the studies of others but also make some critical analysis
- The perspectives of the students are enriched with feedbacks
- During evaluation, students gain empathy
- Students' motivation for learning increases

On the other hand, some limitations of peer education are as follows (Brown, 1998, Magin and Helmore, 2001, Zariski 1996; Van den Berg et. al., 2006):

- Students may not have ability and maturity for evaluation
- Students may not take the evaluation seriously
- Students may have a negative attitude towards peer education
- Students may affect each other during evaluation
- Students may consider it as an additional burden

Validity in Peer Evaluation

The most important difficulty of peer evaluation is to be able to maintain the validity of evaluation at an acceptable level (Dochy et al., 1999). The most important way of ensuring this is to prepare clear and understandable criteria, to make evaluation by hiding the names or to use more than one peer evaluation for one study (Falchikov, 2001). The success of the evaluation is closely related with how much the students adopt criteria and process. It will be useful, if lecturers make student active in this process.

Topping (1998) reviewed studies on peer evaluation, made from 1980 till 1996 about many topics in higher education. He revealed that there was high correlation between lecturer evaluation scores and peer evaluation scores in twenty five studies of thirty one, which he reviewed. This demonstrates that the reviewed peer evaluations have a high validity. Similarly, Falchikov and Goldfinch (2000) reviewed forty eight studies of peer evaluation and found that peer evaluation results show similarity with lecturer evaluation results.

Points to be Paid Attention Concerning Peer Evaluation

Topping (1998) emphasized the importance of peer evaluation to be on time and personal in an effective peer evaluation process. Moreover, it is obvious that more quality peer evaluations will be more effective in learning. For example, Smith et al. (2002) reported that feedbacks as well as markings increase the transparency of the peer evaluation, self-confidence of the student and learning results. Topping (1998) stated that different types of feedbacks also have different effects on student and learning.

Peer evaluation can be made by hiding the names of evaluator and of the person, who is evaluated (Dochy, F., Segers, M. ve Sluijsmans, D. 1999). Thus, it is ensured that friendship relations affect evaluation results at minimum level. Peer evaluation can be made individually or in groups. If the evaluation is made by more than one evaluator, it will be more effective and consistent. The inconsistencies, still exist in evaluation results indicate, for educator, that there are problems in evaluation criteria and show that evaluator has some trouble with fair evaluation.

METHOD

In this chapter, information concerning the model, sample, tools for data collection, analysis and interpretation of the research is included.

Research Model

This study is a relation research, which has been designed for revealing, if similar results are gained when peer evaluation and lecturer evaluation are compared. For this purpose, the relation between peer evaluation figures, gained from students, and scores of lecturers is examined through correlation method. Peer evaluation scores and lecturer evaluation scores constitute the variables of the research.

Sample

This research was made on 48 students, who continue 3rd class at Department of Computer and Instructional Technologies Teaching, Gazi Faculty of Education, Gazi University in 2006-2007 Education Year Spring Term and take "Specific Teaching Methods I" classes.

Evaluation Activity

In this study, students developed web-based learning environments as a project. They first selected one of the strategies introduced in Specific Teaching Methods I class and then selected one topic included in primary school curriculum and after that they combined knowledge on developing story board and Dreamweaver programme, they gained during Authoring Language and Practices in Internet Environment classes. During the first two weeks of the project, students developed their story boards. During the following 3 weeks, they formed the first draft of the project and next week they presented it in the classroom and received feedbacks from peers and lecturers. Then during subsequent 3 weeks, they completed their projects. During the last two weeks, they presented the finalized version of their projects and peer evaluations were made. During this period, students were supported concerning participation in '1st National CITT Students General Assembly' with their projects in order to their share projects, they developed, with other students of Computer and Instructional Technologies Teaching Department and to be able to increase motivation level and students achieved degrees in the mentioned competition and won prizes.

Collecting and Analyzing Data

In this study, data was gained by using Student Peer Evaluation Form, which was prepared in the classroom with students. While developing the form, class discussion was executed and factors affecting the group work and project were taken into consideration. Ideas, gained as the result of discussion were classified and evaluation criteria were finalized.

Students filled the form at the end of term in electronic environment and sent them to the lecturers of the class through e-mail. These forms were gathered in a file in electronic environment. Then the evaluations, made by students and lecturer were transformed into scores in a scale of 100.

Peer evaluation scores and lecturer evaluation scores constitute the variables of the research. Both figures show continuity between one and a hundred. For the purpose of the study, the relation between these two continuous variables was examined by calculating Pearson Correlation Coefficients. It is appropriate to calculate Pearson Correlation Coefficient in order to find and interpret the amount of the relation between these two continuous variables (Büyüköztürk, 2005).

Peer Evaluation Form

Peer evaluation form is composed of two parts. It included 30 criteria, 12 were in the first part and 18 were in the second part.

In the 1st part, students evaluated performances of themselves and peers in term project study of OFD 394 Specific Teaching Methods class. Evaluation scale was composed of four units (0: very bad, 1: bad, 2: good, 3: very good) and included the following criteria:

1. Contribution to group meetings
2. Contribution to group discussions
3. Contribution to group works
4. Contribution to group decisions
5. Contribution to group harmony
6. Bear group responsibility
7. Cooperation and helping
8. Bear individual responsibility
9. Do one's duties
10. Behave one's friends positively
11. Contribution to the whole study
12. Would like to work in another project next time

In the 2nd part, students evaluated project products of other groups and of their own group. Evaluation scale was composed of four units (0: very bad, 1: bad, 2: good, 3: very good) and included the following criteria:

1. Use of technology
2. Design of interface
3. Practicability
4. Originality
5. Attractiveness
6. Appropriateness to the target group

7. Presentation of content
8. The related of instructional content and activities with the real life
9. Appropriateness to the instructional method(s), which were determined
10. Appropriateness to individual learning differences
11. Appropriateness to ethical values
12. Adequacy tools of interaction with lecturer
13. Adequacy of tools of interaction with other students
14. Adequacy of tools of determination of learning deficiencies and feedback
15. Adequacy of tools of self-evaluation
16. Adequacy of tools of source and learning support
17. Adequacy of tools of learning evaluation
18. The studies(s) you most liked

FINDINGS

The score distributions that the students have as a result of peer evaluation and lecturer evaluation are listed from low scores to the high and shown in Table 1 and Table 2 on 1/100 scale.

Table 1. Order of score, gained from peer evaluation.

| | | | | |
|----|----|----|----|-----|
| 43 | 55 | 70 | 78 | 88 |
| 43 | 56 | 71 | 80 | 90 |
| 44 | 60 | 72 | 80 | 90 |
| 44 | 61 | 73 | 83 | 92 |
| 48 | 62 | 73 | 84 | 92 |
| 48 | 62 | 74 | 84 | 96 |
| 49 | 66 | 74 | 86 | 100 |
| 50 | 66 | 75 | 87 | 100 |
| 54 | 66 | 76 | 87 | |
| 54 | 67 | 78 | 88 | |

Table 2. Order of score, gained from lecturer evaluation.

| | | | | |
|----|----|----|----|-----|
| 50 | 62 | 71 | 80 | 90 |
| 50 | 62 | 72 | 81 | 92 |
| 51 | 63 | 72 | 82 | 94 |
| 54 | 64 | 72 | 84 | 94 |
| 56 | 66 | 74 | 84 | 98 |
| 57 | 68 | 75 | 84 | 98 |
| 58 | 68 | 76 | 86 | 100 |
| 58 | 68 | 78 | 86 | 100 |
| 60 | 70 | 78 | 87 | |
| 62 | 70 | 80 | 88 | |

As it is seen in Table 3, the scores of peer evaluation ($\bar{x}=71,22$) are lower than the scores of lecturer evaluation ($\bar{x}=74,43$) by 2,21 scores on average. In both evaluations, the average figure appeared to be 73. However, the most frequently seen figure in peer evaluation is 66, whereas this came out 62 in lecturer evaluation. In lecturer evaluation (SS=14,08) the distribution of scores happened to have number ranges closer to each other, when compared to peer evaluation (SS=16,35). In peer evaluation, it appeared that scores were distributed in 57 score interval while this distribution range is 50 in lecturer evaluation. When the way of distribution is examined, as it is seen in Figure 1 and Figure 2, the scores that the students gained as a result of peer evaluation and lecturer evaluation were distributed in a normal way.

Table 3. Statistics on central tendency and central distribution of peer evaluation and lecturer evaluation score distributions.

| | Peer Evaluation | Lecturer Evaluation |
|--------------------|-----------------|---------------------|
| Number | 48 | 48 |
| Average | 71,22 | 74,43 |
| Median | 73 | 73 |
| Mod | 66 | 62 |
| Standard Deviation | 16,35 | 14,08 |
| Skewness | -,158 | ,087 |
| Kurtosis | -,997 | -,905 |
| Range | 57 | 50 |

Figure 1. Peer Evaluation score distribution histogram graphic.

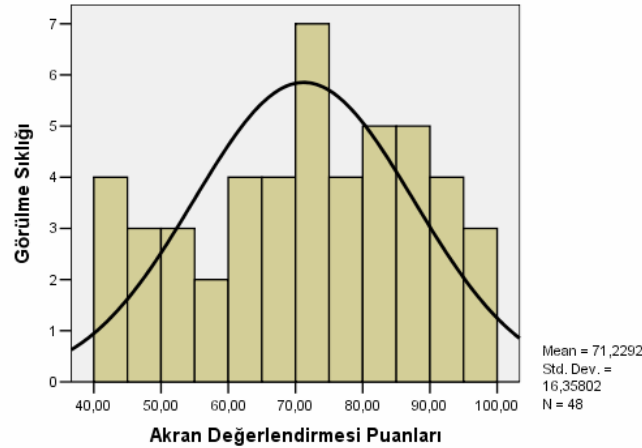
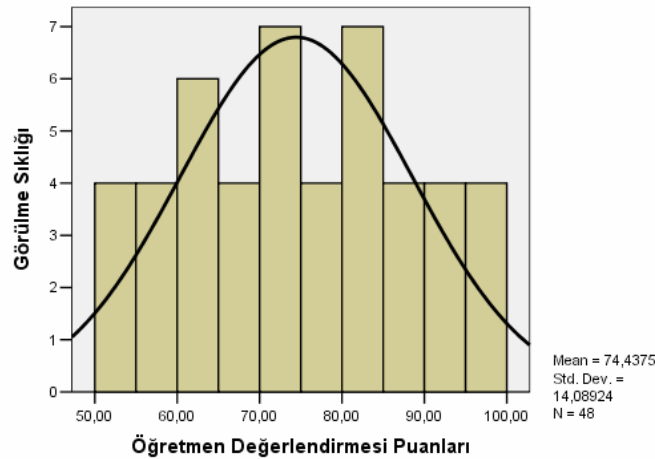


Figure 2. Lecturer Evaluation score distribution histogram graphic.



The relation between peer evaluation and lecturer evaluation was examined by calculating Pearson Correlation values through correlation method. As it is seen Table 4, there is high level relation, both positive and significant, between peer evaluation scores and lecturer evaluation scores ($r=0.991$, $p<.01$). Accordingly, students getting high scores from lecturers also get high scores from peers, and similarly students getting low scores from lecturers also get high scores from peers.

Table 4. The correlation between peer evaluation and lecturer evaluation.

| | | Peer Evaluation | Lecturer Evaluation |
|---------------------|--|-----------------|---------------------|
| Peer Evaluation | Pearson correlation significance (2-directional) | 1 | ,991 ,000 |
| | N | 48 | 48 |
| Lecturer Evaluation | Pearson correlation significance (2-directional) | ,991 ,000 | 1 |
| | N | 48 | 48 |

DISCUSSION OF RESULTS

As a result of the research, when we consider the lecturer evaluations as reference, it can be stated that student peer evaluations show high similarity. This result supports the results of studies conducted earlier by Topping (1998) and Falchikov and Goldvich (2000). Topping (1998) revealed that there was high correlation between lecturer evaluation scores and peer evaluation scores in twenty five studies of thirty one, he reviewed, moreover Falchikov and Goldfinch (2000) examined forty eight peer evaluation studies and found that peer evaluation results show similarity with lecturer evaluation results.

In conclusion, in the light of findings of this research, peer evaluation is suggested to the educators as an alternative evaluation method that can be applied more frequently in higher education. In practice, the importance of preparation of clear and understandable criteria, evaluation made without giving the names or using more than one peer evaluation for one study, should be taken into consideration (Falchikov, 2001). As for the success of the evaluation, it is important how much students adopt criteria and process. It will be useful if lecturers make student active in this process. In this context, criteria can be prepared together with students.

REFERENCES

- Brown, S. (1998), *Peer Assessment in Practice*, Birmingham: SEDA.
- Büyüköztürk, Ş. (2005), *Sosyal bilimler için veri analizi el kitabı. (5. baskı)*, Ankara: Pegem Yayıncılık.
- Dochy, F., Segers, M. & Sluijsmans, D. (1999), "The use of self-, peer and co-assessment in higher education: a review", *Studies in Higher Education*, 24, 331–350.
- Falchikov, N. (1995), "Peer feedback marking: developing peer assessment". *Innovations in Education and Training International*, 32, 175-187.
- Falchikov, N., Goldfinch, J. (2000), "Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks", *Review of Educational Research*, 70:3, 287-322.
- Falchikov, N. (2001), *Learning together; peer tutoring in higher education*, London, Routledge-Falmer.
- Fallows, S. & Chandramohan, B. (2001), "Multiple approaches to assessment: reflections on use of tutor, peer and self assessment", *Teaching in Higher Education*, 6:2, 229-246.
- Freeman, M. (1995), "Peer assessment by groups of group work", *Assessment and Evaluation in Higher Education*, 20, 289–300.
- Fry, S. A. (1990), "Implementation and Evaluation of Peer marking in Higher Education", *Assessment and Evaluation in Higher Education*, 15, 177-189.
- Holroyd, C. (2000), "Are assessors professional? Student assessment and the professionalism of academics", *Active Learning in Higher Education*, 1, 28–44.
- Klein, J.T. (1990), *Interdisciplinarity: history, theory and practice*, Detroit: Wayne State University Press.
- Magin, D. & Helmore, P. (2001), "Peer and teacher assessments of oral presentations: how reliable are they?", *Studies in Higher Education*, 26: 287-298.
- Topping, K. (1998), "Peer assessment between students in colleges and universities", *Review of Educational Research*, 68, 249–276.
- Vygotsky, L.S. (1978), *Mind in Society: The development of higher psychological processes*, Cambridge, MA: Harvard University Press.
- Van den Berg, I., Admiraal, W. & Pilot, A. (2006), "Design principles and outcomes of peer assessment in higher education", *Studies in Higher Education*, 31:3, 341-356.